

Supplementary Data

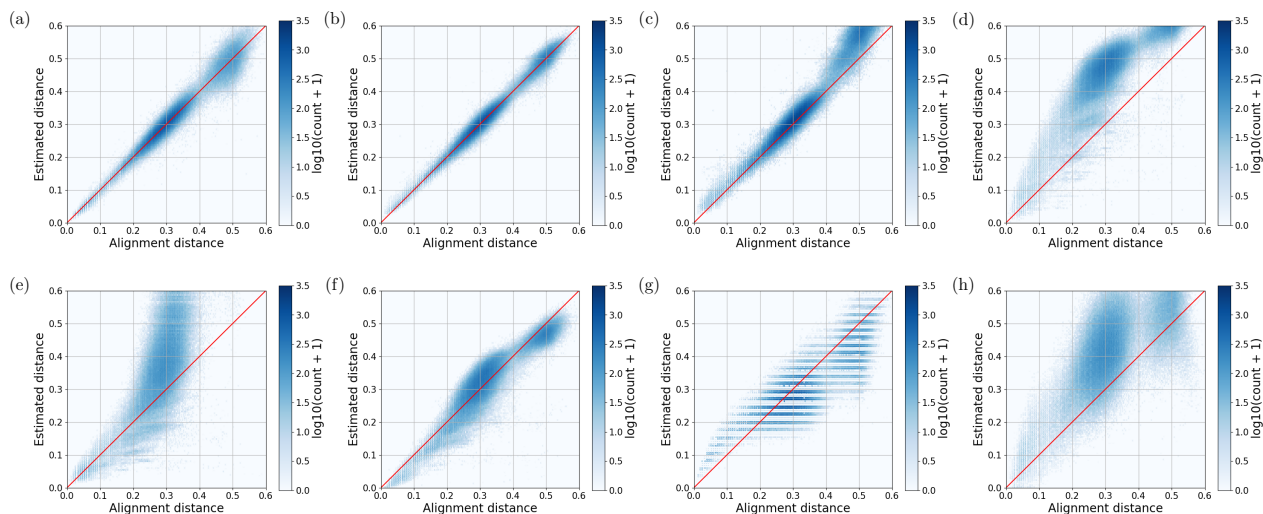


Fig. 1. Visualization of alignment distances versus estimated distances computed by eight methods: (a) AsMac, (b) SENSE, (c) NeuroSEED, (d) ACS, (e) Kr, (f) kmacs ($k=2$), (g) k -mer ($k=3$), and (h) FFP ($l=9$), performed on the Qiita dataset. Each dot represents a sequence pair, and the color of a hex bin represents the number of sequence pairs in the bin. The number in a parenthesis is the parameter that achieved the best result for the corresponding method.

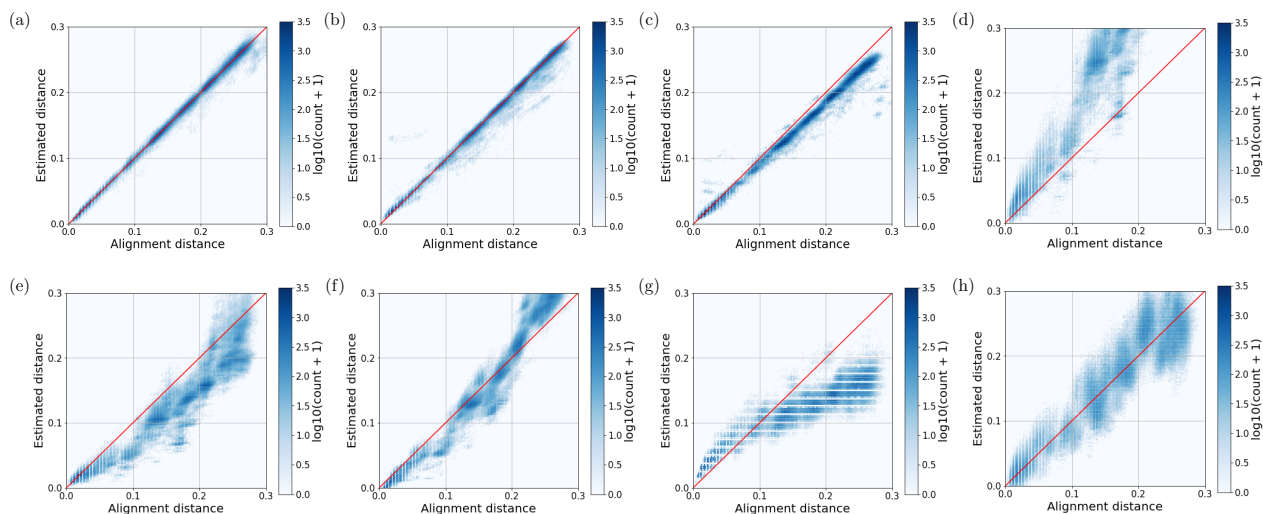


Fig. 2. Visualization of alignment distances versus estimated distances computed by eight methods: (a) AsMac, (b) SENSE, (c) NeuroSEED, (d) ACS, (e) Kr, (f) kmacs ($k=2$), (g) k -mer ($k=3$), and (h) FFP ($l=10$), performed on the RT988 dataset.

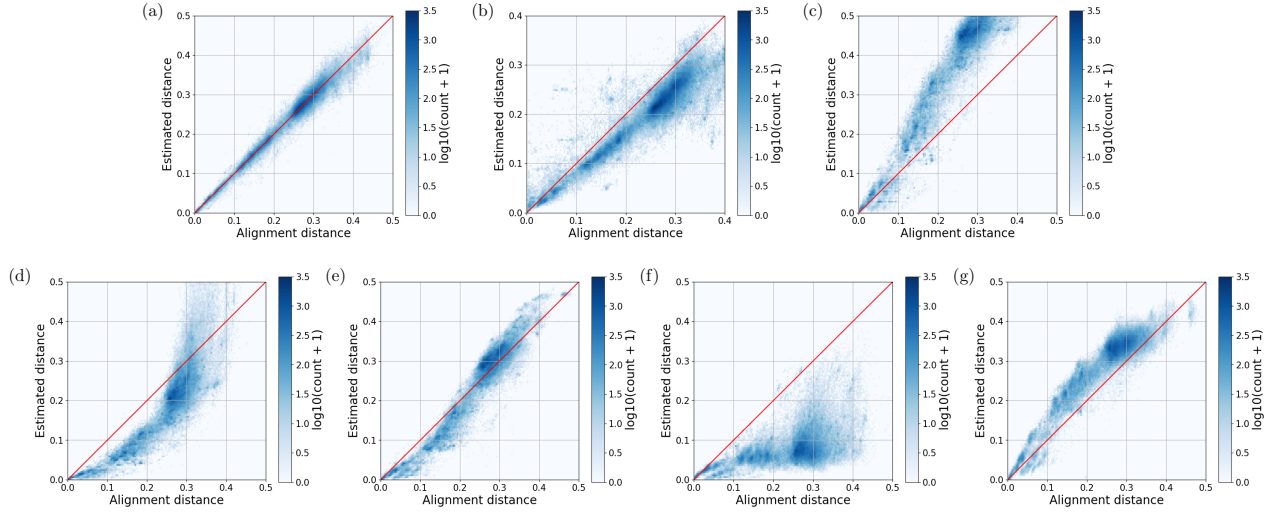


Fig. 3. Visualization of alignment distances versus estimated distances computed by seven methods: (a) AsMac, (b) NeuroSEED, (c) ACS, (d) Kr, (e) kmacs ($k=3$), (f) k -mer ($k=6$), and (g) FFP ($l=13$), performed on the Silva-23S dataset.

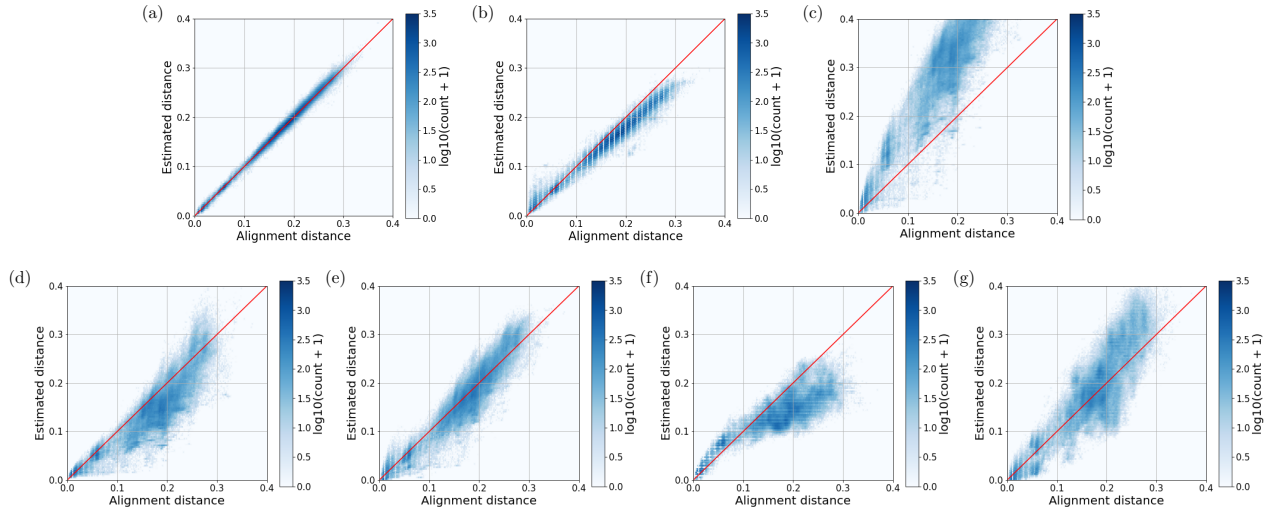


Fig. 4. Visualization of alignment distances versus estimated distances computed by seven methods: (a) AsMac, (b) NeuroSEED, (c) ACS, (d) Kr, (e) kmacs ($k=2$), (f) k -mer ($k=5$), and (g) FFP ($l=12$), performed on the Labonte lake dataset.